(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification[7]:   H04N 7/173

(21) International Application Number: PCT/EP01/02759

(22) International Filing Date: 12 March 2001 (12.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/532,845   21 March 2000 (21.03.2000)   US

(71) Applicant: KONINKLIJKE PHILIPS ELECTRON-ICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors: DIMITROVA, Nevenka; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). AGNIHOTRI, Lalitha; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). MCGEE, Thomas, F.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). MANKOVICH, Nicholas, J.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agent: GRAVENDEEL, Cornelis; Internationaal Octrooibureau B.V., Prof Holstlaan 6, NL-5656 AA Eindhoven (NL).

(81) Designated State (national): JP.

(84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM AND METHOD FOR AUTOMATIC CONTENT ENHANCEMENT OF MULTIMEDIA OUTPUT DEVICE

(57) Abstract: A media display system enhances content by recognizing patterns in the media signal and modifying the media signal responsively to the recognized patterns. For example, in a television broadcast environment, the media signal could be a television program. At one instant, the television could include a logo of a car company. The system would recognize the logo and correlate it with enhanced content stored locally or by using additional input. Based on user preferences (e.g., whether the user is interested in that particular car company) and the correlated enhanced content, the system would modify the broadcast signal in appropriate way. For example, the enhanced content could be a commercial video clip or the phone number of a local car dealer. The modification to the broadcast signal could, in such a case, include overlaying the local car dealer's number on the video signal or buffering the broadcast signal and playing the commercial video clip to perform content enhancement.

System and method for automatic content enhancement of multimedia output device

The present invention relates to a video system that recognizes patterns in digitized images, and more particularly to such systems that isolate symbols or a series of symbols, such as text characters and/or logos in video data streams, and displays information or projects a sound related to the symbols based on a user's preferences. The invention also relates to a system that processes the audio input, such as words, music, or other sounds and responds by displaying information or projecting a sound related to the audio input.

Recognition of text in document images is well known in the art. Recognition of symbols may be based on similar technology. Document scanners and associated optical character recognition (OCR) software are widely available and well understood. However, detection and recognition of text and other symbols in video frames presents unique problems and requires a very different approach than does text in printed documents. Text in printed documents is usually restricted to single-color characters on a uniform background (plain paper) and generally requires only a simple thresholding algorithm to separate the text from the background. By contrast, symbols in scaled-down video images suffer from a variety of noise components, including uncontrolled illumination conditions. Also, the background frequently moves and symbols may be of different color, sizes, orientation, font styles, etc.

BACKGROUND OF THE INVENTION

Real-time broadcast, analog tape, and digital video are a few examples of video sources that provide educational and entertainment value to an observer. These sources can trigger or re-trigger an observer's interest in a particular topic or product. For example, a display of a BMW® logo on a video screen may spark an interest in the observer concerning the performance of BMW® automobiles or an interest in the locations of local and national BMW® authorized dealers.

These interests can be fleeting and soon forgotten by the observer. The observer may be reluctant to pursue his recently triggered interest, especially if pursuit will interrupt his enjoyment of the current programming. For example, a display of a Coca Cola® logo can spark an interest in the number of calories in a 12 ounce can of Coke®, but the interest may not be great enough to motivate the observer to retrieve a can of Coke® to find

this information and satisfy the interest. If not promptly attended to, this interest may be forgotten soon afterwards.

SUMMARY OF THE INVENTION

5   Briefly, a media display system enhances content by recognizing patterns in the media signal and modifying the media signal responsively to the recognized patterns. For example, in a television broadcast environment, the media signal could be a television program. At one instant, the television could include a logo of a car company. The system would recognize the logo and correlate it with enhanced content stored locally. Based on

10   user preferences (e.g., whether the user is interested in that particular car company) and the correlated enhanced content, the system would modify the broadcast signal in appropriate way. For example, the enhanced content could be a commercial video clip or the phone number of a local car dealer. The modification to the broadcast signal could, in such a case, include overlaying the local car dealer's number on the video signal or buffering the

15   broadcast signal and playing the commercial video clip.

          The present invention recognizes patterns in video and/or audio inputs and retrieves and outputs additional information based on the recognized patterns. Patterns may be recognized using any of a variety of known signal processing techniques. The method of the invention classifies patterns, looks up the class identified in the signal in a database, and

20   outputs additional content corresponding to the recognized class along with the current video/audio signal thereby enhancing it. So, for example, if the logo of a car company were recognized in the video stream and found to correspond to a class in the database, the system would locate additional data corresponding to the logo, for example the name and address of a local dealer, and output this data as text superimposed on the video stream. The response

25   of the system may be customized based on the preferences of the user. So, for example, although the logo for the car company may be classifiable using a database of symbol classes employed by the system, the user may not be interested in that car or have some general switch turned off so that the content enhancement does not occur. User-specific preference data can take the form of a separate database to be used in conjunction with a generic

30   classification database that contains classification data for a large number of symbols, including ones the user does not care about. Alternatively, the user's profile data can be used to set up a symbol classification database such that only classes corresponding to interests of the particular user are stored. The latter has the advantage of making it possible for the volume of data stored locally to be minimized.

Pending US Patent Applications 09/370,931, 09/441,943, and 09/441,949 describe methods and devices for classifying symbols, especially text and text blocks, in video streams. The foregoing US patent applications are hereby incorporated by reference in their entirety as if fully set forth herein. The identical method, or any other suitable ones,

5   may be employed to recognize symbols in a video stream for purposes of implementing the instant invention. Recognition of speech is a mature technology area that is continually being refined. Software that can recognize speech is sufficiently developed to recognize various words, especially if their sound is a trademark or has features that are well-known, such as the famous voice of Tony the Tiger associated with a breakfast cereal. The same basic

10   technology used for speech recognition may be applied to the classification of other sounds as well. Thus, the sound of a car accelerating, the sound of a commercial jingle, etc. can also be classified.

Content enhancements can take many forms. For example, an Internet link could be invoked and indicated to the user either as an on-screen token, as a synthetic speech

15   phrase, etc. A Web-TV®-like system may then provide support to allow the user to invoke a link instantly. Alternatively, the system could play a sound clip related to the recognized pattern. The system can operate automatically by immediately displaying the information once it recognizes a symbol or an array of symbols to which the system is programmed to respond. It could also be programmed to respond automatically when it recognizes a word,

20   phrase, or sound. Examples of potential recognizable patterns include a displayed word, a spoken phrase, a logo or a series of musical notes. Displayed information can consist of text information, a sound clip, objects, faces, pictures, or any other information in any media form (sound, visual, etc.) related to the recognized pattern.

The information can be superimposed on the video display so that the observer

25   can continue to receive the programming while receiving the additional stored information. Alternatively, the video data stream could be automatically buffered so that the additional content does not interfere with the user's enjoyment of an on-going sequence. The process of superimposing text on a screen is described in detail in the U.S. patent no. 5,418,576 entitled "Television Receiver with Perceived Contrast reduction in a Predetermined Area of a Picture

30   where Text is .Superimposed." The same techniques can be employed to insert other visual information such as pictures, distortions of existing visual information (e.g., an embossment of the existing visual field on an icon), faces, etc.

The following are examples of the behavior of the inventive system.

1. An observer watches a Mercedes Benz® commercial on a television screen. The system recognizes the Mercedes logo through the video input and displays information about the latest Mercedes automobile by superimposing the information on the television screen.

2. A video input having the text characters "Martin Luther King" on the lower right hand side is received by the system. The system recognizes the characters as being the name of the civil rights leader and plays a sound clip of his "I have a dream" speech.

3. The system receives and processes the audio input to recognize certain words, phrases, or sounds. In one example, processing the trademarked chimes of the Intel® Corporation may trigger the system to retrieve the names of the board members of Intel. The audio input Microsoft®, as recognized through a speech to text converter, can trigger the system to retrieve personal information about William Gates III.

4. Another example is the system may recognizes the face of Einstein and automatically gives a link to a physics Web site.

5. Enhanced content sources may be, for example, Web URLs, web pages, movie databases, stock trading databases, shopping on-line catalogs, museum information, bookstore on-line, dictionaries, or encyclopedias.

The information displayed and the different patterns the system is programmed to recognize can be set according to the user's preference. In certain instances, it may be preferable to limit the number of recognized patterns so that the observer is not bombarded with information about topics that are not interesting to the observer. Such limitation can take the form of limiting the number of responses per set time frame or responding to only certain selected patterns.

In an embodiment, the recognizable patterns and the corresponding stored information are downloaded from the Internet. In this way, the information can be periodically updated by an outside source so that the observer can be supplied with updated information. The system could be programmed to request responses from the user to the enhanced content and to update the profile accordingly. Alternatively, the system could be programmed to allow the user to set parameters explicitly during a setup procedure.

The system may also incorporate a software switch for turning off this function or a logon feature so that the system can operate according to the preferences of a currently logged-on user. The user can also select the type of information displayed for each patterned input when more than one set of information is available for a single recognized pattern. For example, if the system has both location information and stock price information

for the recognizable pattern for Tiffany's, the system can display one or both of the available information depending on the user's preferences.

Besides the benefit of providing the observer with useful information, the system can also benefit advertisers by directing advertising only to interested users. Thus, the directed advertising information can include special offers on products or services related to the recognized pattern. The delivery of content by advertisers to the set-top boxes of the user for delivery through the system can be administered through the Internet. Users can update their set-top boxes with new content dynamically. A system service provider could allow dynamic updating by advertisers of content and some of the rules for display.

The local content can be controlled dynamically according to the current season (e.g., barbecue grill content might not be stored locally during the winter and snow plowing services might not be stored locally in the summer), according to the advertising campaign underway by the advertiser, etc. as well as according to the local preferences of the user.

Video or audio inputs can also be modified or intentionally programmed to trigger a response by the system. For example, the video/audio input provider can provide programming that accents the recognizable patterns, such as by placing the intended video pattern on the screen with a white background. The enhanced item could be identified with, and function as, a clickable link, say to a locally stored media item, an online-store connection, etc. This feature could be made sensitive to the type of show or some other parameter (e.g., time of day, channel, etc.) so that it would not occur at all times. So, for example, if a person in a talk show were wearing a wardrobe identifiable with a particular designer, outlet stores where the designer's goods could be purchased might be displayed or otherwise made available through a link.

The invention will be described in connection with certain preferred embodiments, with reference to the following illustrative figures so that it may be more fully understood.

With reference to the figures, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this regard, no attempt is made to show structural details of the invention in more detail that is necessary for a fundamental understanding of the invention, the

description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

BRIEF DESCRIPTION OF THE DRAWINGS

5   FIG. 1 is a diagram illustrating the components that may be used to practice the invention.

FIG. 2 is a block diagram of the functional elements that may be used to implement the invention according to one embodiment thereof.

FIG. 3 is a flowchart showing a basic content enhancement method according 10   to an embodiment of the invention.

FIG. 4 is a figurative image of a display showing two recognizable symbols which may trigger a content enhancement.

FIG. 5 is a figurative image of a display showing a website superimposed over a video image illustrating a way of enhancing content without interrupting an ongoing 15   multimedia display.

FIG. 6 is a figurative image of a display showing a website in a picture-in-picture display superimposed over a video image illustrating another way of enhancing content without interrupting an ongoing multimedia display.

20   DETAILED DESCRIPTION OF THE DRAWINGS

Referring to FIG. 1, the invention may be used in connection with the environment of a television with Internet capability. In the embodiment of FIG. 1, a computer 240 sends program information to a television 230. The computer 240 may be equipped to receive the video signal 270 and control the channel-changing function and to 25   provide Internet browser capability. Commands may be entered into the computer 240 via a memory card or disk 220, a remote controller 210 (connected via an IR port 215) or a keyboard 212 or downloaded via network connection. A data link 260 provides Internet connection and an antenna, cable, or satellite link 270 provides audio and/or video data. This could be a telephone line connectable to an Internet service provider or some other suitable 30   data connection. Note that the data and audio/video links 260 and 270 could include the same physical channel. The computer 240 preferably has a mass storage device 235, for example a hard disk, to store program schedule information, program applications and upgrades, and other information. Information about the user's preferences and other data can

WO 01/72040

PCT/EP01/02759

be uploaded into the computer 240 via removable media such as the memory card or disk 220.

Note that many substitutions are possible in the above example hardware environment and all can be used in connection with the invention. The computer can be a
5   set-top box with processing capability. The mass storage can be replaced by volatile memory or non-volatile memory. The data can be stored locally or remotely. In fact, the entire computer 240 could be replaced with a server operating offsite through a link. Rather than using a remote control 210 or keyboard 212 to send commands to the computer 240, these controllers could send commands through a data channel 260 which could be separate from,
10   or the same as, the physical channel carrying the video. The video 270 or other content can be carried by a cable, satellite, RF, or any other physical channel or obtained from a mass storage or removable storage medium. It could be carried by a switched physical channel such as a phone line or a virtually switched channel such as ATM or other network suitable for synchronous data communication. Content could be asynchronous and tolerant of
15   dropouts so that present-day IP networks could be used. Further, the content of the line through which programming content is received could be audio, chat conversation data, web sites, or any other kind of content for which a variety of selections are possible. Data can be received through channels other than the separate data link 260. For example, data can be received through the same physical channel as the video or other content. It could even be
20   provided through removable data storage media such as memory card or disk 220. The remote control 210 can be replaced by a keyboard, voice command interface, 3D-mouse, joystick, or any other suitable input device. Selections can be made by moving a highlighting indicator, identifying a selection symbolically (e.g., by a name or number), or making selections in batch form through a data transmission or via removable media. In the latter
25   case, one or more selections may be stored in some form and transmitted to the computer 240, bypassing the display 170 altogether. For example, batch data could come from a portable storage device (e.g. a personal digital assistant, memory card, or smart card, or downloaded). Such a device could have many preferences stored on it for use in various environments so as to customize the computer equipment to be used.

30   In the embodiment of FIG. 1, an advertiser user interface in the form of an advertiser client process 170 provides data to a Host server 175. The host server sends data to the computer 240 which stores this data selectively on a local data store, for example, a disk 235. The advertiser client process 170 could be implemented via a browser session in which an advertiser, wishing to provide content enhancement through the directed advertising

channel provided by an embodiment of the present invention, could upload data including media content and various control data. The uploaded data is stored on a service host server 175 for control and periodic updating of the viewer system 200.

Referring to FIG. 2, video, audio, and/or other media data are supplied by

5   some source or sources 310, to an output device 350. The media signal is modified, displaced, and/or stored by a hard disk video recorder, DVD-RW or WebTV box or Media Output Combiner / Switch / Buffer / Client 390. A set-top box (not shown), functionally coterminous with computer 240, may provide the latter functionality. A symbol classifier 330 receives the media data and parses the signal to search for recognizable elements. These

10   elements can be graphic images, images, audio sequences, voice fingerprints, or any other classifiable signal. The symbol classifier 330 outputs class identifiers to an enhanced content processor 360. A user profile data store 320 stores user preferences with regard to the enhanced media content that may be displayed. For example, the user profile data may contain an indication that the user is interested in sports cars and that the user does not mind

15   interruptions of broadcast media to receive enhanced content from advertisers relating sports cars. The enhanced content processor 360 applies the class identifier from the symbol classifier to a class / enhanced content correlation data store 370 to find a pointer to media content contained in an enhanced content data store 340. The class identifier and the pointer to the media content contained in the enhanced content data store 340 are combined with the

20   user profile data from user profile data store 320 to determine if a content enhancement should be made. The latter determination is made by the enhanced content processor 360, which generates control parameters indicating the content and instructions for combining the enhanced content with the media data to be applied to an enhanced content output controller 395. The enhanced content output controller 395 in turn takes enhanced content from the

25   enhanced content data store 340, generates instructions for the media output combiner / switch 390 responsively to the instructions from the enhanced content processor and commands from an input device 355. The enhanced content output controller 395 then outputs selected media content and instructions to the media output combiner / switch 390 to modify the media data stream before it is displayed on the output device 350. The user input

30   device 355 also allows the user profile data to be updated in user profile data store 320.

The symbol classifier may utilize any suitable mechanism for classifying the media data signal stream. For example, the methods described in pending US Patent Applications 09/370,931, 09/441,943, and 09/441,949 for classifying symbols, especially text and text blocks, in video streams, may be used and preferred, particularly the neural network

method and system described in the latter application. Recognition of speech is a mature technology area that is continually being refined. Software that can recognize speech is sufficiently developed to recognize various words. The same and related signal processing technology, for example voice-print technology that can be used to identify the voices of particular individuals, may be used by symbol classifier 330. The classes can be trademark sounds or sounds with features that are well-known, such as the famous voice of Tony the Tiger associated with a breakfast cereal. Classes can be defined for sounds like the sound of a car accelerating, the sound of a commercial jingle, etc.

The user profile data store 320 may contain any of a variety of user-modifiable parameters that informs the control processes used in the embodiment described with respect to FIG. 2. The profile data may include any of the following or any other suitable parameters.

1) Enhancement technique for various types of enhanced content such as websites, commercials, text or audio clips, etc., for example, the profile may indicate whether

    a) the media data should be buffered and the display changed to invoke a web site corresponding to classified media element;

    b) a video image should be ghosted and continued in the background (See discussion with reference to FIG. 5) while a website display is shown on top of it;

    c) a link should be placed on the display which can then be selected (for example, using a pointer and button on remote 210);

    d) a picture-in-picture display (See discussion with reference to FIG. 6) may be shown with the additional content such as a commercial, a website, etc.

    e) a text overlay is preferred over a high bandwidth item such as a commercial or infomercial

2) Storage options to allow enhanced content to be book-marked for future display.

3) How classified items should be identified, such as by applying a solarize filter to a portion of the video display and halting (and buffering it), increasing contrast, or switching to an alternate view.

4) The type of content that is of interest to the user, for example, sports cars, beer, weddings, business, technology, literature, weather, etc.

    a) This could include different levels of interest. So, for example, if the user is generally not particularly interested in weather per se, the user could

indicate an interest in receiving enhanced content only if a weather
advisory were issued for the user's locality.

5) What data sources are available for extracting additional user profile data, for example, a
set-top box used for television viewing with enhanced electronic program guide
information may store user preferences with respect to genre, time of day, preferred
channels and programs, etc. that may be used to provide data to the user profile data store
320.

The class / enhanced content correlation data store 370 may be a lookup table
data indicating a correspondence between recognized classes and enhanced content that may
be output. The class / enhanced content correlation data store 370 may also contain data
downloaded from the service host server 175 originating from the advertiser client process
170 indicating certain specific instructions with regard to that content such as an expiration
date for a contest, weather conditions that should obtain before the content is output (e.g.,
only advertise snow plowing services when it is snowing), etc. The data stored by the class /
enhanced content correlation data store 370, thus points to particular items in the enhanced
content data store 340.

The enhanced content processor 360 takes its instructions from user profile
data and the class identifiers supplied by the symbol classifier 330. The enhanced content
processor obtains the vector(s) required to find the relevant content data and supplies this and
control information to the enhanced content output controller 395. Referring to FIGS. 2
through 6, the media data are sampled and processed by the symbol classifier 330 in step A-
1. The symbol classifier attempts, in step A-2, to classify the whole or portions of the media
data until it identifies and classifies a pattern. If a pattern is successfully classified in step A-
3, a class identifier is applied to the enhanced content processor 360. Then, in step A-4, the
enhanced content processor 360 applies any conditions or rules in the user profile data and
the content provider (e.g., advertiser) data stored in class / enhanced content correlation data
store 370 to determine if and what precise content should be used to enhance the media data
stream. If content enhancement is indicated, in step A-5, control passes to step A-6 in which
enhanced content output controller 395 controls media output combiner / switch / buffer /
client 390 to enhance the media signal accordingly. So, for example, if a car company logo
310 is identified in the media data by the symbol classifier 330, a class identifier would be
applied to the enhanced content processor 360. The result might be, depending on rules
obtained from the user profile data store 320 and class / enhanced content correlation data
370, a website overlay 325. Alternatively, if a weather warning message 315 were displayed,

marquis-style, across the screen, a picture-in-picture (PIP) window 330 indicating the availability of a weather-information website or another broadcast channel could be displayed with an invitation to the viewer to switch to the additional content.

Note that, as described above, the concept of enhancing content can include

5   the addition of an icon, selection of which could provide additional media content, or it could simply be the addition of content on the original media stream. An example of the latter would be superimposed text with the phone number of a local Ford(R) dealer in response to a Ford(R) commercial or logo. Thus, the idea of enhanced content can include interactive elements through which the timing, content, scope, etc. of the enhanced content can be

10   controlled by the user in response to the classification of a media element. Thus, for example, the user can decide whether to link to the additional weather information by selecting an ephemeral link (the PIP window 330) or to ignore it. If the user selects the link, more media content is displayed than if the user ignores it. Additionally, if the link is to a live website, not all the media content is supplied through enhanced content data store 340.

15   The enhanced content output controller 395 is responsive to the control parameters applied by the enhanced content processor 360. The control parameters can include a basic process to be followed by the enhanced content output controller 395 along with a set of pointers to specific items of content within the enhanced content data store 340. For example, the basic process might be to highlight a specific region of the display for a

20   specified amount of time, to define a selectable region on the display and to trigger a web link upon selection of that region via the input device 355. The media content could be a URL, a video clip, a sound, or a bit of formatted text. The content along with detailed instructions for modifying the media data stream are sent to the media output combiner / switch / buffer / client 390. The latter implements the instructions to overlay text, buffer the

25   media data, provide Internet client services, to invoke a website, etc. according to the commands from the enhanced content output controller 395. Note that the enhanced content output controller 395 may transmit more than one set of instructions and media items. For example, if the initial phase of enhanced content is the generation of a web link, only the link media data and instructions defining the superimposed link would be transmitted. Upon

30   receipt of a selection via input device 355, further content, such as a URL would then be supplied output to the media output combiner / switch / buffer / client 390.

Note that a variety of input and output devices may be used to implement the current invention. The input device 355 could be a light pen, a TV-style remote controller, a keyboard, a mouse, or any other suitable device for generating commands. The output device

350 could be a TV, a monitor, a TV-wall, broadcast or multicast channel, chat client, Internet console, or any other suitable device. Media data can include any type of information content including text, video, audio, live data by closed circuit system, chat data, IP packets, etc. The output device 350 could also include more than one physical device. For example,

5   it could be multiple monitors where one is used to output enhanced content and the other outputs the original media data.

Note that other data 332 may be used by the enhanced content processor 360 to make decisions. For example, electronic program guide data may indicate the genre of the media data stream (e.g., a TV broadcast in the comedy genre). Other data that may be

10   relevant to decisions to supply enhanced content may include the time of day, day of year, current weather, name of the broadcast item appearing in the media signal, etc.

Note that the processing of video or other streaming information could be implemented as a back end process rather than at the client (e.g., the remote terminal or television set-top box near the viewer). In such an implementation, only control information

15   need be transmitted to the client process and the processing and storage capacity could be reduced. That is, all the possible symbol classes and information for classifying raw data could be stored at the back end processor. Just the data required to produce the content enhancement would be transmitted. The content enhancement data could be transmitted as embedded control information using any suitable process of such as video watermarks, data

20   inserted in the blanking interval, etc. Control data could also be delivered by XML or other meta standards for multimedia data packaging including MPEG-7, ATSC, DVB, etc.

It is evident to those skilled in the art that the invention is not limited to the details of the foregoing illustrative embodiments, and that the present invention may be embodied in other specific forms without departing from the spirit or essential attributes

25   thereof. The present embodiments are therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention be indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

CLAIMS:

1.   A method of enhancing the content of a video output comprising the steps of:
storing enhanced content data on a server (175);
downloading said enhanced content data to a client system (240, 175);
classifying portions of a media data stream connected to said client system;
5   modifying said media data stream responsively to profile data and said enhanced content data
to produce enhanced media data; and
outputting said enhanced media data stream.

2.   A method as in claim 1, wherein said step of classifying includes recognizing
10   a graphical pattern (310, 315).

3.   A method as in claim 1, wherein said step of modifying includes the addition
of one or more of at least one of a visual or audio element corresponding to a selectable link
to a website (325).

15

4.   A method as in claim 1, wherein said step of classifying includes recognizing
speech.

5.   A system for enhancing the content of a broadcast media stream, comprising:
20   an output device (230);
a signal modification device (240, 175) connected to apply a media data signal to said output
device and to receive said broadcast data stream;
a user profile data store (175, 235) storing user profile data;
a media content data store (175, 235) storing media content items;
25   a controller, connected to said user profile and media content data stores, programmed to
recognize a portion of said broadcast data stream;
said controller being programmed to control said signal modification device responsively to
said user profile data and said media content items.

6.   A system as in claim 5, wherein said broadcast media stream includes a video
stream, said media content items include text, and said controller is programmed to control
said signal modification device to overlay an image corresponding to said text on said video
stream.

5

7.   A system as in claim 5, wherein said output device is a television (230).

8.   A system as in claim 5, wherein said broadcast media stream is a television
broadcast signal and said portions include portions of a television signal that are ordinarily
10   displayed as visible elements.

9.   A system as in claim 5, wherein said output device includes at least two
separate video displays and said media content items are output on at least one of said at least
two separate video displays and said broadcast data stream on the other of said two separate
15   video output devices.

10.   A method of enhancing the content of a video output comprising the steps of:
classifying portions of a media data stream deliverable to a client system;
classifying features in said media data stream;
20   generating control signals responsively to said step of classifying;
modifying said media data stream responsively to said control signals and profile data
defining preferences of a class of user;
outputting a result of said step of modifying.

25   11.   A method as in claim 10, wherein at least one of said control signals are
generated at a server (175) and delivered to said client embedded in said media data stream.

12.   A device for displaying media content on an output device (230), comprising:
a pattern classifier connected to receive a media broadcast signal and to output class
30   identifiers responsively to patterns recognized in said media broadcast signal;
a media content data store (175, 235) containing said media content;
a user preference data store (175, 235) holding user preference data;
a controller (240, 175) programmed to output, on said output device, selected portions of said
media content responsively to said user profile and said class identifiers.

13.   A device as in claim 12, wherein said controller is programmed to combine said selected portions of said media with said broadcast signal to generate a combined output signal to said output device.
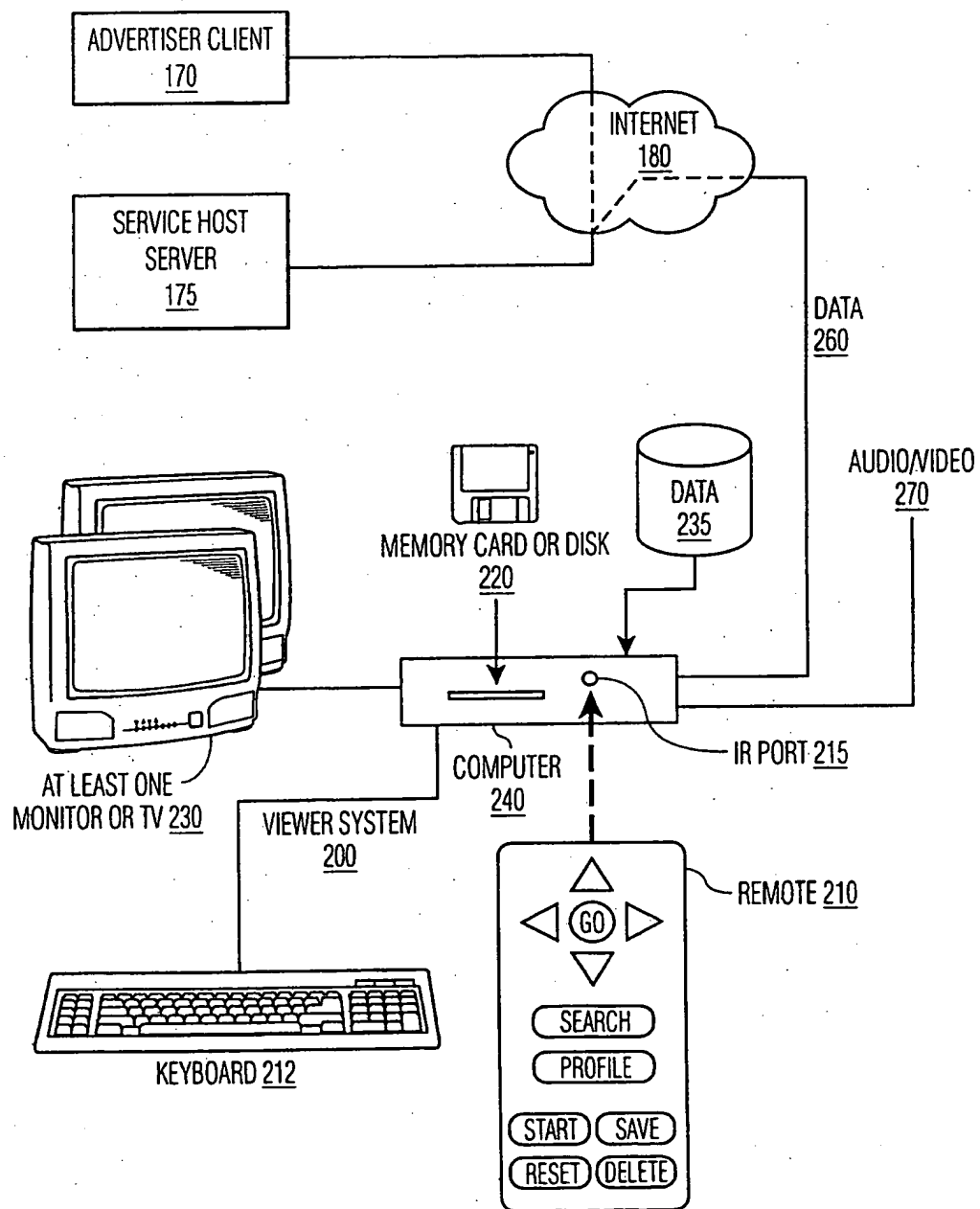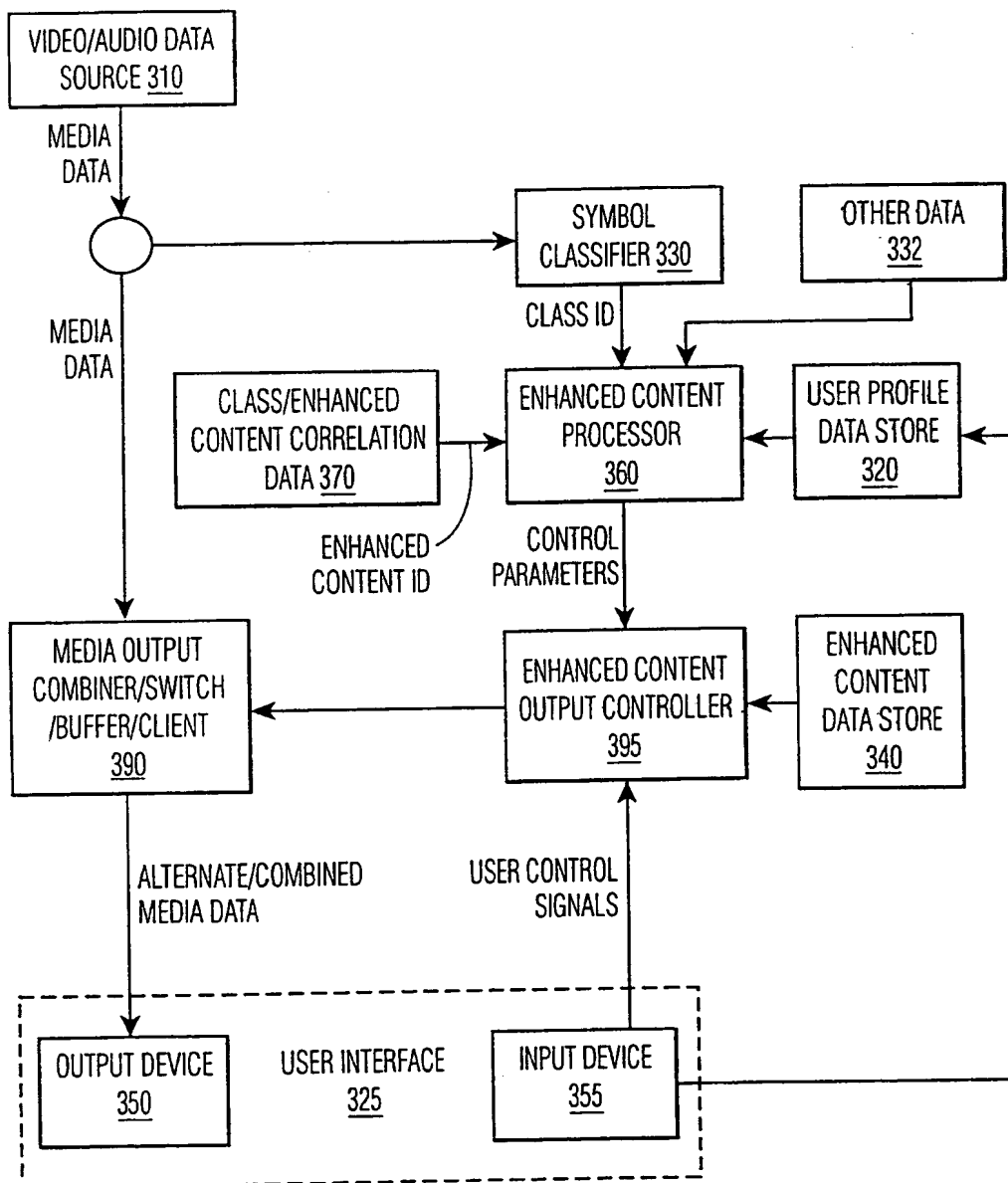
FIG. 1

FIG. 2

```
        ┌─────────────────┐
        │       A-1       │
        │   CAPTURE AND   │◄──────────────────────────○◄──────┐
        │ PREPROCESS VIDEO│                                    │
        │  AUDIO INPUTS   │                                    │
        └─────────────────┘                                    │
                 │                                             │
                 ▼                                             │
        ┌─────────────────┐                                    │
        │       A-2       │                                    │
        │ ATTEMPT PATTERN │                                    │
        │  CLASSIFICATION │                                    │
        └─────────────────┘                                    │
                 │                                             │
                 ▼                                             │
              ╱  A-3  ╲          NO                            │
             ╱ ANY PATTERNS╲──────────────────────►○          │
             ╲ CLASSIFIABLE?╱                       │          │
              ╲           ╱                         │          │
                 │ YES                              │          │
                 ▼                                  │          │
        ┌─────────────────┐                         │          │
        │       A-4       │                         │          │
        │FIND CLASS(ES) IN LOOKUP                    │          │
        │ TABLE AND APPLY USER                       │          │
        │PROFILE AND OTHER RULES                     │          │
        └─────────────────┘                         │          │
                 │                                  │          │
                 ▼                                  │          │
              ╱  A-5  ╲          NO                 │          │
             ╱SHOULD ENHANCED╲────────────────────────────────┘
             ╲  CONTENT BE  ╱                       │
              ╲ APPLIED? ╱                          │
                 │ YES                              │
                 ▼                                  │
        ┌─────────────────┐                         │
        │       A-6       │                         │
        │GENERATE CORRESPONDING───────────────────────────────┘
        │CONTENT ENHANCEMENT                        │
        └─────────────────┘
```
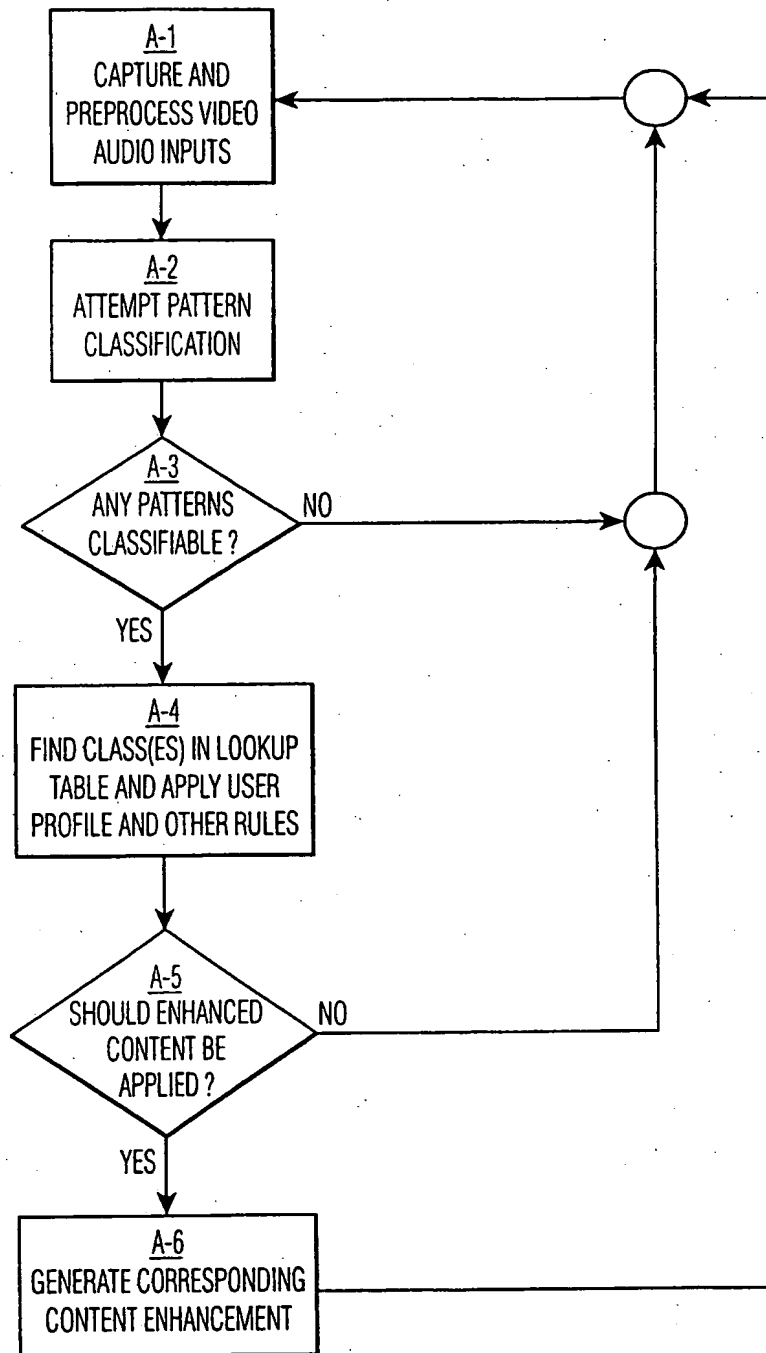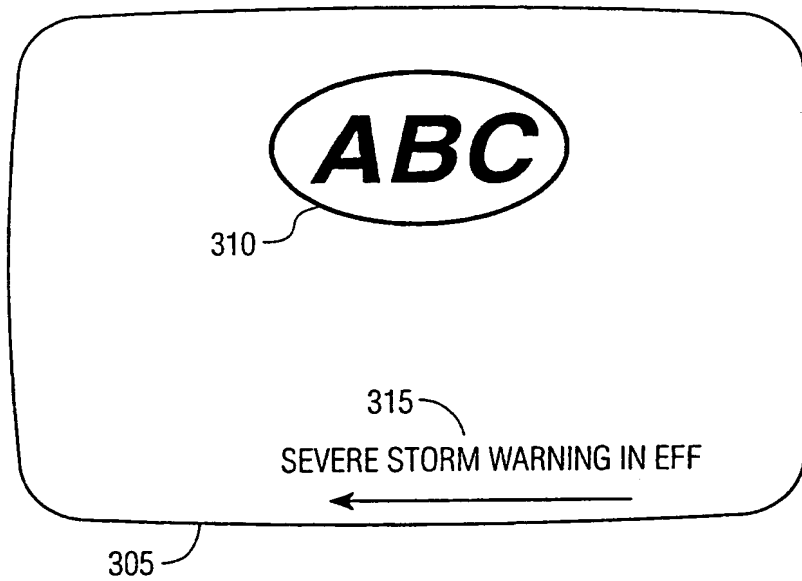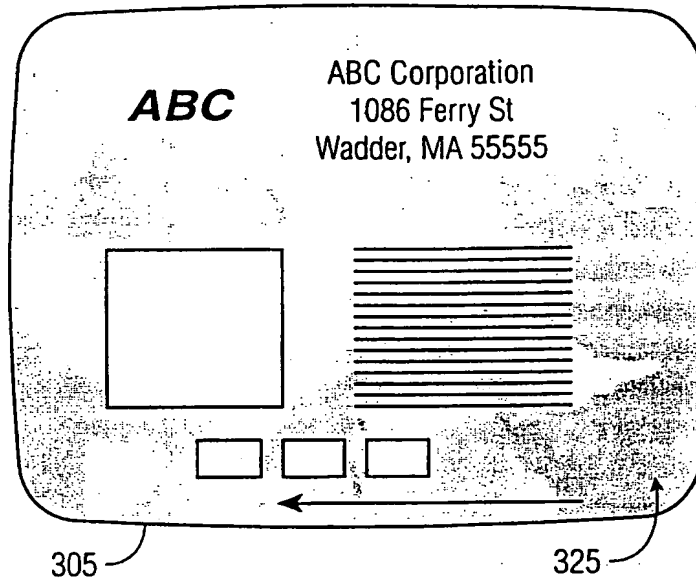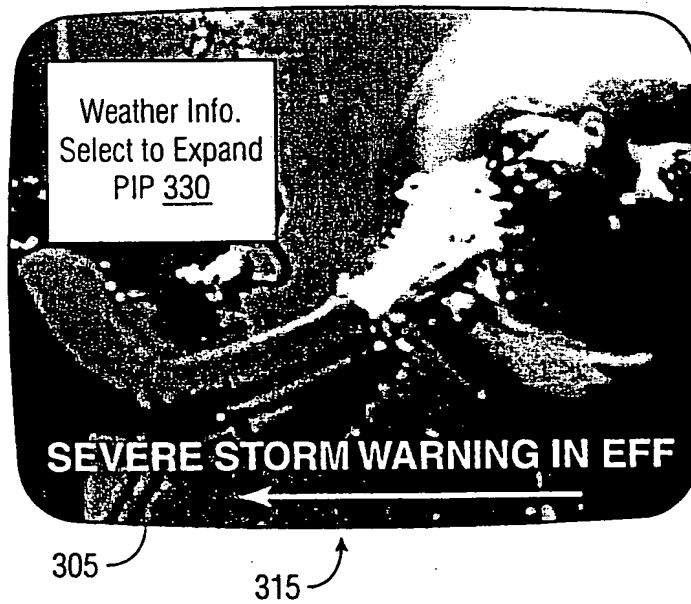
FIG. 3

FIG. 4

FIG. 5

FIG. 6

THIS PAGE BLANK (USPTO)